IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor:	SPIVAK et al.	Examiner:	Adam L. Basehoar
Application No.:	09/815,591	Art Unit:	2178
Filed:	March 23, 2001	Docket No.:	EMCCP074
Title:	METHOD AND APPARATUS FOR GENERATING METADATA FOR A DOCUMENT		

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in a prepaid envelope addressed to: Mail Stop Appeal Brief – Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on:

7/9, 2009. Elaine Nguyen

Elaine Nguyen

REPLY BRIEF PURSUANT TO 37 C.F.R. §41.41

MAIL STOP APPEAL BRIEF - PATENTS
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

This Reply Brief is in response to the Examiner's Answer mailed on May 12, 2009.

I. REAL PARTY IN INTEREST

The real party of interest in the present application is EMC Corporation.

II. RELATED APPEALS AND INTERFERENCES

PURSUANT TO 37 C.F.R. §41.37(c)(1)(ii), Appellant hereby notifies the Board of Patent Appeals that Appellant, the Appellant's Legal Representative, and the Assignee do not

know of any appeals or interferences that will directly affect or be directly affected by or have any bearing on the Board's decision in the pending appeal.

III. STATUS OF CLAIMS

Claims 1, 4-8, 10-16, and 18-21 are currently pending in the application, and are attached hereto as an appendix. Claims 2, 3, 9, and 17 have been canceled. Claims 1, 4-8, 10-16, and 18-21 were finally rejected by the Examiner and are the subject this appeal.

IV. STATUS OF AMENDMENTS

No amendment has been filed since the Final Office Action. All previously submitted amendments are believed to have been entered

V. SUMMARY OF CLAIMED SUBJECT MATTER

The claimed subject matter of Claim 1 relates to a method of processing a document. Figure 2 illustrates an example of the method for processing a document. In the example shown in Figure 2, a document is converted to a common format (e.g., 204 of Figure 2, lines 24-25 of page 7 of the specification), and concepts are recognized using a document modeling module (e.g., 208 of Figure 2, lines 28-30 of page 7 of the specification). Figure 4 illustrates an example of a document modeling module, which is discussed without limitation in line 4 of page 11 through line 7 on page 13. In the discussion of the document modeling module and the process of Figure 5 illustrating a process for the document modeling module (e.g., lines 23 of page 13 of the specification through line 30 page 16 of the specification), it is indicated that concepts represent a basic idea that is expressed in the document (e.g., lines 18-21 of page 11 of the specification) and that recognizing concepts includes: identifying a set of features (e.g., line 18 of page 11 of the specification) where the features represent evidence of a concept (e.g., line 19 of page 11 of the specification); calculating concept weights using feature weights (e.g., lines 8-9 of page 15 of the specification), where concept weights represent a recognition confidence level for a concept (e.g., line 5-6 of page 15 of the specification) and where the recognition confidence

level is calculated for each paragraph (e.g., lines 28-31 of page 15 of the specification); and comparing the concept weight to a threshold value (e.g., line 19-20 of page 16 of the specification). Figure 7 further illustrates an example of a document modeling module, which is discussed without limitation in line 16 of page 18 of the specification through line 3 of page 26 of the specification. In the further discussion of the document modeling module, it is indicated that recognized concepts are associated using a concept model (e.g., lines 10-15 of page 20 of the specification); indicating a concept type associated with the document (e.g., lines 13-14 of page 20 of the specification), where concept type comprises a group of concepts representing a similar idea (e.g., lines 1-2 of page 19 of the specification); and identifying that the document is responsive to a search query using the concept type (e.g., lines 5-9 of page 20 of the specification).

The claimed subject matter of Claim 4 relates to a computer-readable medium to direct a computer to function in a specified manner comprising instructions for processing a document. Figure 1 illustrates an example of a computer readable medium storing instructions for processing a document (e.g., lines 11-13 of page 5 of the specification) including computer programs implementing methods for processing. Figure 2 illustrates an example of a method for processing a document. In the example shown in Figure 2, a document is converted to a common format (e.g., 204 of Figure 2, lines 24-25 of page 7 of the specification), and concepts are recognized using a document modeling module (e.g., 208 of Figure 2, lines 28-30 of page 7 of the specification). Figure 4 illustrates an example of a document modeling module, which is discussed without limitation in line 4 of page 11 through line 7 on page 13. In the discussion of the document modeling module and the process of Figure 5 illustrating a process for the document modeling module (e.g., lines 23 of page 13 of the specification through line 30 page 16 of the specification), it is indicated that concepts represent a basic idea that is expressed in the document (e.g., lines 18-21 of page 11 of the specification) and that recognizing concepts includes: identifying a set of features (e.g., line 18 of page 11 of the specification) where the features represent evidence of a concept (e.g., line 19 of page 11 of the specification); calculating concept weights using feature weights (e.g., lines 8-9 of page 15 of the specification), where concept weights represent a recognition confidence level for a concept (e.g., line 5-6 of page 15 of the specification) and where the recognition confidence level is calculated for each paragraph (e.g., lines 28-31 of page 15 of the specification); and comparing the concept weight to a

threshold value (e.g., line 19-20 of page 16 of the specification). Figure 7 further illustrates an example of a document modeling module, which is discussed without limitation in line 16 of page 18 of the specification through line 3 of page 26 of the specification. In the further discussion of the document modeling module, it is indicated that recognized concepts are associated using a concept model (e.g., lines 10-15 of page 20 of the specification); indicating a concept type associated with the document (e.g., lines 13-14 of page 20 of the specification), where concept type comprises a group of concepts representing a similar idea (e.g., lines 1-2 of page 19 of the specification); and identifying that the document is responsive to a search query using the concept type (e.g., lines 5-9 of page 20 of the specification). The document modeling module may additionally assign one or more auto-attributes (e.g., lines 21-22 of page 22 of the specification, also line 24 of page 22 through line 18 of page 24 of the specification).

The claimed subject matter of Claim 7 relates to a computer-readable medium to direct a computer to function in a specified manner comprising instructions for processing a document. Figure 1 illustrates an example of a computer readable medium storing instructions for processing a document (e.g., lines 11-13 of page 5 of the specification) including computer programs implementing methods for processing. Figure 2 illustrates an example of a method for processing a document. In the example shown in Figure 2, a document is converted to a common format (e.g., 204 of Figure 2, lines 24-25 of page 7 of the specification), and concepts are recognized using a document modeling module (e.g., 208 of Figure 2, lines 28-30 of page 7 of the specification). Figure 4 illustrates an example of a document modeling module, which is discussed without limitation in line 4 of page 11 through line 7 on page 13. In the discussion of the document modeling module and the process of Figure 5 illustrating a process for the document modeling module (e.g., lines 23 of page 13 of the specification through line 30 page 16 of the specification), it is indicated that concepts represent a basic idea that is expressed in the document (e.g., lines 18-21 of page 11 of the specification) and that recognizing concepts includes: identifying a set of features (e.g., line 18 of page 11 of the specification) where the features represent evidence of a concept (e.g., line 19 of page 11 of the specification); calculating concept weights using feature weights (e.g., lines 8-9 of page 15 of the specification), where concept weights represent a recognition confidence level for a concept (e.g., line 5-6 of page 15 of the specification) and where the recognition confidence level is calculated for each paragraph (e.g., lines 28-31 of page 15 of the specification); and comparing the concept weight to a

threshold value (e.g., line 19-20 of page 16 of the specification). Figure 7 further illustrates an example of a document modeling module, which is discussed without limitation in line 16 of page 18 of the specification through line 3 of page 26 of the specification. In the further discussion of the document modeling module, it is indicated that recognized concepts are associated using a concept model (e.g., lines 10-15 of page 20 of the specification); indicating a concept type associated with the document (e.g., lines 13-14 of page 20 of the specification), where concept type comprises a group of concepts representing a similar idea (e.g., lines 1-2 of page 19 of the specification); and identifying that the document is responsive to a search query using the concept type (e.g., lines 5-9 of page 20 of the specification).

The claimed subject matter of Claim 14 relates to a computer for processing a document. Figure 1 illustrates an example of a computer for processing a document (e.g., lines 11-13 of page 5 of the specification) including a processor (e.g., 116 of Figure 1); a memory (e.g., 118 of Figure 1); and an interface (e.g., 114 or 138 of Figure 1). The memory includes stored instructions for modules for processing a document including a document modeling module (e.g., 122 of Figure 1). The memory stores computer programs, including methods, for processing a document. Figure 2 illustrates an example of a method for processing a document. In the example shown in Figure 2, a document is converted to a common format (e.g., 204 of Figure 2, lines 24-25 of page 7 of the specification), and concepts are recognized using a document modeling module (e.g., 208 of Figure 2, lines 28-30 of page 7 of the specification). Figure 4 illustrates an example of a document modeling module, which is discussed without limitation in line 4 of page 11 through line 7 on page 13. In the discussion of the document modeling module and the process of Figure 5 illustrating a process for the document modeling module (e.g., lines 23 of page 13 of the specification through line 30 page 16 of the specification), it is indicated that concepts represent a basic idea that is expressed in the document (e.g., lines 18-21 of page 11 of the specification) and that recognizing concepts includes: identifying a set of features (e.g., line 18 of page 11 of the specification) where the features represent evidence of a concept (e.g., line 19 of page 11 of the specification); calculating concept weights using feature weights (e.g., lines 8-9 of page 15 of the specification), where concept weights represent a recognition confidence level for a concept (e.g., line 5-6 of page 15 of the specification) and where the recognition confidence level is calculated for each paragraph (e.g., lines 28-31 of page 15 of the specification); and comparing the concept weight to a

threshold value (e.g., line 19-20 of page 16 of the specification). Figure 7 further illustrates an example of a document modeling module, which is discussed without limitation in line 16 of page 18 of the specification through line 3 of page 26 of the specification. In the further discussion of the document modeling module, it is indicated that recognized concepts are associated using a concept model (e.g., lines 10-15 of page 20 of the specification); indicating a concept type associated with the document (e.g., lines 13-14 of page 20 of the specification), where concept type comprises a group of concepts representing a similar idea (e.g., lines 1-2 of page 19 of the specification); and identifying that the document is responsive to a search query using the concept type (e.g., lines 5-9 of page 20 of the specification).

The claimed subject matter of Claim 19 relates to a computer-readable medium to direct a computer to function in a specified manner comprising instructions for processing a document. Figure 1 illustrates an example of a computer readable medium storing instructions for processing a document (e.g., lines 11-13 of page 5 of the specification) including computer programs implementing methods for processing. Figure 2 illustrates an example of a method for processing a document. In the example shown in Figure 2, a document is converted to a common format (e.g., 204 of Figure 2, lines 24-25 of page 7 of the specification), and concepts are recognized using a document modeling module (e.g., 208 of Figure 2, lines 28-30 of page 7 of the specification). Figure 4 illustrates an example of a document modeling module, which is discussed without limitation in line 4 of page 11 through line 7 on page 13. In the discussion of the document modeling module and the process of Figure 5 illustrating a process for the document modeling module (e.g., lines 23 of page 13 of the specification through line 30 page 16 of the specification), it is indicated that concepts represent a basic idea that is expressed in the document (e.g., lines 18-21 of page 11 of the specification) and that recognizing concepts includes: identifying a set of features (e.g., line 18 of page 11 of the specification) where the features represent evidence of a concept (e.g., line 19 of page 11 of the specification); calculating concept weights using feature weights (e.g., lines 8-9 of page 15 of the specification), where concept weights represent a recognition confidence level for a concept (e.g., line 5-6 of page 15 of the specification) and where the recognition confidence level is calculated for each paragraph (e.g., lines 28-31 of page 15 of the specification); and comparing the concept weight to a threshold value (e.g., line 19-20 of page 16 of the specification). Figure 7 further illustrates an example of a document modeling module, which is discussed without limitation in line 16 of

page 18 of the specification through line 3 of page 26 of the specification. In the further discussion of the document modeling module, it is indicated that recognized concepts are associated using a concept model (e.g., lines 10-15 of page 20 of the specification); indicating a concept type associated with the document (e.g., lines 13-14 of page 20 of the specification), where concept type comprises a group of concepts representing a similar idea (e.g., lines 1-2 of page 19 of the specification); and identifying that the document is responsive to a search query using the concept type (e.g., lines 5-9 of page 20 of the specification). The document modeling module may additionally assign one or more auto-attributes (e.g., lines 21-22 of page 22 of the specification, also line 24 of page 22 through line 18 of page 24 of the specification).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Claims 1, 4-5, 7-8, 10-16, and 18-21 stand rejected under 35 U.S.C. 103(a) as being unpatentable over Doerre et. al. (U.S. Patent No. 6,446,061) in view of Weiser et. al. (U.S. Patent No. 5,982,507) in further view of Chakrabarti (U.S. Patent No. 6,418,433) in further view of Russell-Falla (U.S. Patent No. 6,675,162) in further view of McKeown (U.S. Patent No. 6,473,730).

Claim 6 stands rejected under 35 U.S.C. 103(a) as being unpatentable over Doerre et. al. (U.S. Patent No. 6,446,061) in view of Weiser et. al. (U.S. Patent No. 5,982,507) in further view of Chakrabarti (U.S. Patent No. 6,418,433) in further view of Russell-Falla (U.S. Patent No. 6,675,162) in further view of McKeown (U.S. Patent No. 6,473,730) in further view of W3C's, Extensible Markup Language (XML) 1.0", 02/10/98, pp1-2, http://www.w3.org/TR/1998/REC-xml-19980210.

VII. ARGUMENT

CLAIMS 1, 4-8, 10-16, and 18-21 REJECTED UNDER 35 U.S.C. 103(a)

For the reasons set forth below, Appellant respectfully submits that the Examiner has erred in maintaining the 35 U.S.C 103(a) rejection of Claims 1, 4-8. 10-16, and 18-21 as being unpatentable over Doerre, Weiser, Chakrabarti, Russell-Falla, and McKeown because the Final

Office Action does not establish prima facie obviousness of a claimed invention, because all the claim features are not taught or suggested in the prior art.

To establish prima facie obviousness of a claimed invention, all the claim features must be taught or suggested in the prior art. See in. re. Royka, 490 F.2d 981, 180 USPQ 580 (CCPA 1974). Further, all words in a claim must be considered in judging the patentability of that claim against the prior art. *In re Wilson*, 424 F.2d 1382, 1385, 165 USPQ 494, 496 (CCPA 1970). And, also if an independent claim is nonobvious under 35 U.S.C. 103, then any claim depending therefrom is nonobvious. *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988).

The Examiner acknowledges in the Final Office Action that "Doerre does not teach receiving a search query associated with said concept type and identifying based at least in part on the association of the concept type with said document, that said document is responsive to said search query" (p. 4, line 10, Final Office Action). Chakrabarti also does not teach receiving a search query associated with a concept type and identifying, based at least in part on the association of the concept type with a document, that the document is responsive to a search query. Specifically, claims 1, 7, and 14 recite "identifying, based at least in part on the association of the concept type with said document, that said document is responsive to said search query." In fact, Chakrabarti teaches that "a user can search the database 30 efficiently for Web pages of interest, i.e., for only Web pages relating to the topic on which database 30 is focused" (Col. 5, lines 22-25) and that the database "is focused only on the topic such that the system includes no topically comprehensive database of the World Wide Web." (Col. 3, lines 47-50). Chakrabarti does not teach that *identifying, based at least in part on the association of the concept type with the document, that the document is responsive to the search query*, but rather "in response to the search query, accessing a crawl database containing only information pertaining to Web pages related to a limited number of predefined topics." (Col. 3, lines 54-56) For Chakrabarti, the database Web pages are segregated into the database prior to a search of the database. So, even in the event that a concept type were the same as the predefined topics, Chakrabarti does not identify based at least in part on the association of the concept type with the document, that the document is responsive to the search query. Further, the predefined topics stem from a "relatively narrow topic or range of topics that are of interest to a particular source" and are not "a comprehensive, universal set of all topics on the Web" (col. 6, lines 1-4), which

are not the same as a concept type because the specification defines that "a concept type groups one or more concepts that represents similar ideas." (page 19, lines 1-2 of the specification). Further, Chakrabarti teaches:

> "It will readily be appreciated that the database 30 does not contain a listing of Web pages unrelated to the predefined topic, but only relevant Web pages. Consequently, the browser 58 will quickly respond to the query when the query is related to the predefined topic; otherwise, for queries unrelated to the topic, no response will be available from the database 30. Under such a circumstance, the browser 58 can access a conventional crawler database." (col. 6, lines 40-48).

So, no association of the search query is made with the predefined topics, in fact rather the database returns no response and another database is queried.

For the reasons above, the combination of Doerre and Chakrabarti does not establish prima facie obviousness of a claimed invention. It is therefore believed that claims 1, 7, and 14 are allowable. Claims 4-6, 8 & 10-13, and 15-16 & 18-21 depend from claims 1, 7, and 14 and are believed to be allowable for the same reasons as above.

For the reasons set forth below, Appellant respectfully submits that the Examiner has erred in maintaining the 35 U.S.C 103(a) rejection of Claims 1, 4-8. 10-16, and 18-21 as being unpatentable over Doerre, Weiser, Chakrabarti, Russell-Falla, and McKeown because the Final Office Action does not set forth a legally sufficient suggestion or motivation to combine or modify Doerre's method of generating a content taxonomy of a multitude of documents with Weiser, Chakrabarti, or Russell-Falla.

"When a rejection depends on a combination of prior art references, there must be some teaching, suggestion, or motivation to combine the references." See, e.g., *In re Rouffet,* 149 F.3d 1350, 1355, 47 USPQ2d 1453, 1456 (Fed. Cir. 1998). While the Supreme Court has adopted a flexible approach and reasoned against applying the teaching, suggestion, motivation test in a rigid manner, the Court continues to recognize that when the prior art teaches away from combining certain known elements, discovery of a successful means of combining them is more likely to be nonobvious. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398 (2007). If the proposed

modification would render the prior art invention being modified unsatisfactory for its intended purpose, then there is no suggestion or motivation to make the proposed modification. *In re Gordon*, 733 F.2d 900 (Fed. Cir. 1984). Further, if the proposed modification or combination of the prior art would change the principle of operation of the prior art invention being modified, then the teachings of the references are not sufficient to render the claims *prima facie* obvious. *In re Ratti*, 270 F.2d 810 (CCPA 1959). Rejections on obviousness grounds cannot be sustained by mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398 (2007); *In re Kahn*, 441 F.3d 977, 988 (Fed. Cir. 2006).

Specifically, it would not have been obvious to one of ordinary skill in the art at the time of invention to rationally combine Doerre and Weiser. Weiser teaches integrating headerless apparatus, such as scanners and copiers, into a messaging environment (e.g., email environment). In contrast, Doerre teaches classifying batches of documents in one or more existing databases to generate a taxonomy for the documents. Adding headers to documents is very different from categorizing batches into multiple topics. Therefore, in addition to not being rational to combine, there is a basic conflict between the principles of operation of Doerre and Weiser.

Also, it would not have been obvious to one of ordinary skill in the art at the time of invention to rationally combine Doerre and Chakrabarti. Chakrabarti teaches building topic specific databases to provide databases for searching. In contrast, Doerre teaches classifying documents in one or more existing databases to generate a taxonomy for the documents. Preselecting items for a database is different from categorizing items in a database. Therefore, in addition to not being rational to combine, there is a basic conflict between the principles of operation of Doerre and Chakrabarti.

Further, it would not have been obvious to one of ordinary skill in the art at the time of invention to rationally combine Doerre and Russell-Falla. Russell-Falla teaches identifying instances of a category of information (e.g., pornography) in a real-time stream of information (e.g., viewed web pages) in order to be able to control viewing of the category of information. In contrast, Doerre teaches classifying batches of documents in one or more existing databases to generate a taxonomy for the documents. Real-time screening items for a single topic is different

from categorizing batches into multiple topics. Therefore, in addition to not being rational to combine, there is a basic conflict between the principles of operation of Doerre and Russell-Falla.

It is therefore believed that claims 1, 7, and 14 are allowable. Claims 4-6, 8 & 10-13, and 15-16 & 18-21 depend from claims 1, 7, and 14 and are believed to be allowable for the same reasons as above.

For the reasons set forth below, Appellant respectfully submits that the Examiner has erred in maintaining the 35 U.S.C 103(a) rejection of Claims 4 and 19 as being unpatentable over Doerre, Weiser, Chakrabarti, Russell-Falla, and McKeown because the Final Office Action does not establish prima facie obviousness of a claimed invention, because all the claim features are not taught or suggested in the prior art.

With regard to claims 4 and 19, Doerre teaches categorization by subject and that "the classification schema can be discovered automatically through clustering techniques (the machine finds the groups or clusters and assigns to each cluster a generalized title or cluster label that becomes the class name)." (Col. 2 lines 24-27) In contrast, an auto-attribute can assign a predetermined label based on, for example, rules and that labels can include labels of interest such as "Useful Document." (Line 30 page 22 through line 20 page 23 of the specification). Generating rule-based attributes automatically is different from labeling or naming subject groups or clusters.

It is therefore believed that claims 4 and 19 are allowable.

Response to Examiner's Response to Argument

The Examiner has mailed a response to Appellant's arguments on 05/12/2009. Appellant respectfully addresses the Examiner's response with the following:

In regard to independent claims 1, 7, and 14, the Appellant respectfully maintains that Doerre and Chakrabarti fail to teach or suggest "receiving a search query associated with a concept type and identifying, based at least in part on the association of the concept type with a document, that the document is responsive to a search query". As the Examiner agrees, Doerre

does not teach "receiving a search query associated with a concept type and identifying, based at least in part on the association of the concept type with a document, that the document is responsive to a search query." The Examiner indicates that Chakrabarti generates a database with a list of web pages that are associated with a predefined topic/concept. However, a concept type, which "groups one or more concepts that represents similar ideas" (page 19, lines 1-2 of the specification) and which is associated with "a concept association for the two or more recognized concepts," is not associated with a search query or with a document according to Chakrabarti. The applicant's specification teaches: "rather than merely identifying documents that express one or more concepts of interest, the conceptual taxonomy 800 [or the conceptual model] enables a search engine 130 to identify one or more documents by specifying one or more concept types of interest" (page 20, lines 7-9). Applicants claims recite that a "concept type [is] associated with said document using the conceptual model" and that the "search query [is] associated with said concept type" and that "based at least in part on the association of the concept type with said document, [identifying] that said document is responsive to said search query." Note further that the conceptual model "includes the concept association for the two or more recognized concepts." This is not the same as a search query receiving a response of a predefined topic's list of web pages.

VIII. CLAIMS APPENDIX

Listing of Claims:

1. (Previously Presented) A computer-implemented method of processing a document, said method comprising:

converting a document into a common format document;

recognizing two or more concepts in said common format document, wherein said two or more concepts each represent a basic idea expressed in said common format document, wherein recognizing said two or more concepts includes for each of said two or more concepts:

identifying a plurality of features in said common format document, wherein said plurality of features represents evidence of one of said two or more concepts in said common format document;

calculating a concept weight for one of said two or more concepts using a plurality of feature weights associated with said plurality of features, wherein said concept weight represents a recognition confidence level for one of said two or more concepts, and wherein the recognition confidence level for one of said two or more concepts is calculated for each paragraph in the common format document; and

comparing said concept weight with a predetermined threshold value;

recognizing a concept association for the two or more recognized concepts associated with a conceptual model that includes the concept association for the two or more recognized concepts;

indicating a concept type associated with said document using the conceptual model, wherein the concept type comprises a group of one or more concepts that represent a similar idea;

receiving a search query associated with said concept type; and

identifying, based at least in part on the association of the concept type with said document, that said document is responsive to said search query.

2. (Canceled)

3. (Canceled)

4. (Previously Presented) The computer-implemented method of claim 1, further comprising:

based on said conceptual model, generating an auto-attribute, said auto-attribute being a descriptive label for said common format document.

5. (Previously Presented) The computer-implemented method of claim 1, further comprising:

based on said conceptual model, assigning said common format document to a subject category.

6. (Original) The computer-implemented method of claim 1, wherein said converting includes converting said document into a common format document that is in an XML format.

7. (Previously Presented) A computer-readable medium to direct a computer to function in a specified manner, comprising instructions for:

converting a document into a common format document;

recognizing two or more concepts in said common format document, wherein said two or more concepts each represent a basic idea expressed in said common format document, wherein recognizing said two or more concepts includes for each of said two or more concepts:

identifying a plurality of features in said common format document, wherein said plurality of features represents evidence of one of said two or more concepts in said common format document;

calculating a concept weight for one of said two or more concepts using a plurality of feature weights associated with said plurality of features, wherein said concept weight represents a recognition confidence level for one of said two or more concepts, and wherein the recognition confidence level for one of said two or more concepts is calculated for each paragraph in the common format document; and

comparing said concept weight with a predetermined threshold value;

recognizing a concept association for the two or more recognized concepts associated with a conceptual model that includes the concept association for the two or more recognized concepts;

indicating a concept type associated with said document using the conceptual model, wherein the concept type comprises a group of one or more concepts that represent a similar idea;

receiving a search query associated with said concept type; and

identifying, based at least in part on the association of the concept type with said document, that said document is responsive to said search query.

8. (Previously Presented) The computer-implemented method of claim 7, wherein the conceptual model includes a concept dictionary.

9. (Canceled)

10. (Previously Presented) The computer-implemented method of claim 7, further comprising incorporating said two or more concepts into said conceptual model in the event that the recognition confidence level exceeds the predetermined threshold value.

11. (Previously Presented) The computer-implemented method of claim 7, wherein the conceptual model includes a noise dictionary.

12. (Previously Presented) The computer-implemented method of claim 7, further comprising: assigning a subject category to said document based at least in part upon said conceptual model.

13. (Previously Presented) The computer-implemented method of claim 12, wherein assigning the subject category follows an auto-categorization rule.

14. (Previously Presented) A computer, comprising:

a processor; and

a memory connected to said processor, wherein said memory includes:

a document modeling module, said document modeling module having:

a first module configured to direct said processor to recognize two or more

concepts in a document, wherein each of said two or more concepts represents a basic idea expressed in said document, wherein recognizing said two or more concepts includes for each of said two or more concepts:

identifying a plurality of features in said document, wherein said plurality of features represents evidence of one of said two or more concepts in said document;

calculating a concept weight for one of said two or more concepts using a plurality of feature weights associated with said plurality of features, wherein said concept weight represents a recognition confidence level for one of said two or more concepts, and wherein the recognition confidence level for one of said two or more concepts is calculated for each paragraph in the document; and

comparing said concept weight with a predetermined threshold value;

; and

a second module configured to recognize a concept association for the two or more recognized concepts associated with a conceptual model that includes the concept association for the two or more recognized concepts; and

a third module configured to indicate a concept type associated with said document using the conceptual model, wherein the concept type comprises a group of one or more concepts that represent a similar idea; and

an interface configured to receive a search query, wherein when said search query is associated with said concept type, said document is identified as being responsive to said search query based at least in part on the association of the concept type with said document.

15. (previously presented) The computer of claim 14, wherein said memory further includes:

a document integration module, said document integration module having:

a twelfth module configured to direct said processor to convert said document to a common format.

16. (Previously Presented) The computer of claim 15, wherein said document integration module further has:

a fourth module configured to direct said processor to separate a text portion from said document; and

a fifth module configured to direct said processor to incorporate said text portion in said document in the common format.

17. (Canceled)

18. (Original) The computer of claim 14, wherein said memory further includes:

a modeling directory,

and wherein said document modeling module further has:

a ninth module configured to direct said processor to store said conceptual model in said modeling directory.

19. (Original) The computer of claim 14, wherein said document modeling module further has:

a tenth module configured to direct said processor to generate an auto-attribute based upon said conceptual model, wherein said auto-attribute represents a descriptive label for said document.

20. (Original) The computer of claim 14, wherein said document modeling module further has:

an eleventh module configured to direct said processor to categorize said document in a category of a plurality of categories based upon said conceptual model.

21. (Previously Presented) The computer-implemented method of claim 1, wherein the conceptual model includes a concept association dictionary.

IX. EVIDENCE APPENDIX

Not Applicable.

X. RELATED PROCEEDINGS APPENDIX

Not Applicable.

Respectfully submitted,
VAN PELT, YI & JAMES LLP

Terrence L. Wong
Registration No. 55,741

10050 N. Foothill Blvd., Suite 200
Cupertino, CA 95014
Telephone: 408-973-2578